01-06-00

CHOATE, HALL & STEWART
A PARTNERSHIP INCLUDING PROFESSIONAL CORPORATIONS
EXCHANGE PLACE
53 STATE STREET
BOSTON, MASSACHUSETTS 02109-2891
TELEPHONE (617) 248-5000
FACSIMILE (617) 248-4000

January 5, 2000

EXPRESS MAIL NO.EK044517371US

Assistant Commissioner for Patents
Washington, DC 20231
Box Patent Application

TRANSMITTAL LETTER

Dear Sir:

Please find enclosed a patent application as follows:

Applicant(s): Kevin A. Jarrell, Matthew D. Shair
Title: COMBINATORIAL BIOLOGY

No. Pages Specification: 17; No. Pages Claims: 2; No. Pages of Drawings 5; No. Pages Abstract:1; Unexecuted Forms: Declaration 1; Appointment of Attorney; Establishing Right of Assignee to Take Action (37 CFR 3.73(b)) and Return Postcards; 1

Basic Fee: $380.00

Additional Fees:

Total Number of Claims in excess of 20:
Number of independent claims minus 3 times:
Multiple dependent claims ($130):

Total Filing Fee $380.00

Enclosed please find a check in the amount of $380.00 to cover the total filing fee. Please charge any additional fees that are required or credit any overpayments to our Deposit Account No. 03-1721.

If this application is found otherwise to be INCOMPLETE, or if at any time it appears that a TELEPHONE CONFERENCE with counsel would helpfully advance prosecution, please telephone the undersigned at any time.

Kindly acknowledge receipt of the foregoing application by returning the self-addressed postcard.

Respectfully submitted,

Brenda Herschbach Jarrell, Reg. No. 39,223

ATTORNEY DOCKET NO.: 0342941-0037

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant: Jarrell et al.
Serial No.:
Filed: January 5, 2000
For: COMBINATORIAL BIOLOGY

VERIFIED STATEMENT (DECLARATION)
CLAIMING SMALL ENTITY STATUS
37 CFR 1.9(f) and 1.27(d)
NON-PROFIT ORGANIZATION

I hereby declare that I am an official empowered to act on behalf of the non-profit organization identified below:

NAME OF ORGANIZATION: Trustees of Boston University
ADDRESS OF ORGANIZATION: 148 Baystate Road, Boston, Massachusetts 02215
TYPE OF ORGANIZATION:

(X) University or other organization of higher education

( ) Non-profit scientific or educational organization under a statute of a state of the United States of America
NAME OF STATE: _____
CITATION OF STATUTE: _____

( ) Would qualify as tax exempt under the Internal Revenue Service Code (26 USC 501(a) and 501(c) (3) if the organization were located in the United States of America

( ) Tax exempt under the Internal Revenue Service Code (26 USC 501(a) and 501(c) (3).

( ) Would qualify as a non-profit scientific or educational organization under a statute of a state of the United States of America
NAME OF STATE: _____
CITATION OF STATUTE: _____

I hereby declare that the above-identified non-profit organization qualifies as a non-profit organization as defined in 37 CFR 1.9(e), for purposes of paying reduced fees under section 41(a) and (b) of Title 35, United States Code.

I hereby declare that rights under contract or law have been conveyed to and remain with the above-identified non-profit organization with regard to the invention titled:

TITLE: Combinatorial Biology

by:

INVENTORS: Kevin A. Jarrell, Matthew D. Shair

described in:

(X) the specification filed herewith

( ) U.S. Patent Application Serial Number _____
filed _____

( ) U.S. Patent Number _____
issued _____

If the rights of the above-identified non-profit organization are not exclusive, each individual, concern, or organization having rights to the invention is listed below* and no rights to the invention are held by any person, other than the inventor(s), who could not qualify as a small business concern under 37 CFR 1.9(d) or a non-profit organization under 37 CFR 1.9(e).

*NOTE: Separate verified statements are required from each named person, concern, or organization having rights to the invention, averring to their status as small entities. (37 CFR 1.27).

FULL NAME: President and Fellows of Harvard College

ADDRESS: 17 Quincy Street, Cambridge, MA 02139

( ) INDIVIDUAL ( ) SMALL BUSINESS CONCERN (X) NON-PROFIT ORGANIZATION

I acknowledge the duty to file, in this application or patent, notification of any change in status resulting in loss of entitlement to small entity status prior to paying, or at the time of paying, the earliest of the issue fee or any maintenance fee due after the date on which status as a small entity is no longer appropriate. (37 CFR 1.28(b))

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under §1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application, any patent issuing thereon, or any patent to which this verified statement is directed.

NAME OF PERSON SIGNING: Matthew J. Burns

TITLE IN ORGANIZATION: Assistant Treasurer

ADDRESS OF PERSON SIGNING: Trustees of Boston University
108 Baystate Road, Boston, MA 02215

SIGNATURE: _____ DATE: _____

ATTORNEY DOCKET NO.: 0342941-0037

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant: Jarrell et al.
Serial No.:
Filed: January 5, 2000
For: COMBINATORIAL BIOLOGY

VERIFIED STATEMENT (DECLARATION)
CLAIMING SMALL ENTITY STATUS
37 CFR 1.9(f) and 1.27(d)
NON-PROFIT ORGANIZATION

I hereby declare that I am an official empowered to act on behalf of the non-profit organization identified below:

NAME OF ORGANIZATION: President and Fellows of Harvard College
ADDRESS OF ORGANIZATION: 17 Quincy Street, Cambridge, MA 02139
TYPE OF ORGANIZATION:

(X) University or other organization of higher education

( ) Non-profit scientific or educational organization under a statute of a state of the United States of America
NAME OF STATE: _____
CITATION OF STATUTE: _____

( ) Would qualify as tax exempt under the Internal Revenue Service Code (26 USC 501(a) and 501(c) (3) if the organization were located in the United States of America

( ) Tax exempt under the Internal Revenue Service Code (26 USC 501(a) and 501(c) (3).

( ) Would qualify as a non-profit scientific or educational organization under a statute of a state of the United States of America
NAME OF STATE: _____
CITATION OF STATUTE: _____

I hereby declare that the above-identified non-profit organization qualifies as a non-profit organization as defined in 37 CFR 1.9(e), for purposes of paying reduced fees under section 41(a) and (b) of Title 35, United States Code.

I hereby declare that rights under contract or law have been conveyed to and remain with the above-identified non-profit organization with regard to the invention titled:

TITLE: Combinatorial Biology
by:
INVENTORS: Kevin A. Jarrell, Matthew D. Shair
described in:

(X) the specification filed herewith
( ) U.S. Patent Application Serial Number _____
filed _____
( ) U.S. Patent Number _____
issued _____

If the rights of the above-identified non-profit organization are not exclusive, each individual, concern, or organization having rights to the invention is listed below* and no rights to the invention are held by any person, other than the inventor(s), who could not qualify as a small business concern under 37 CFR 1.9(d) or a non-profit organization under 37 CFR 1.9(e).

*NOTE: Separate verified statements are required from each named person, concern, or organization having rights to the invention, averring to their status as small entities. (37 CFR 1.27).

FULL NAME: Trustees of Boston University
ADDRESS: 148 Baystate Road, Boston, Massachusetts 02215
( ) INDIVIDUAL ( ) SMALL BUSINESS CONCERN (X) NON-PROFIT ORGANIZATION

I acknowledge the duty to file, in this application or patent, notification of any change in status resulting in loss of entitlement to small entity status prior to paying, or at the time of paying, the earliest of the issue fee or any maintenance fee due after the date on which status as a small entity is no longer appropriate. (37 CFR 1.28(b))

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under §1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application, any patent issuing thereon, or any patent to which this verified statement is directed.

NAME OF PERSON SIGNING: Joyce Brinton
TITLE IN ORGANIZATION: Director of the Office of Technology and Licensing
ADDRESS OF PERSON SIGNING: Harvard University
Technology and Trademark Licensing
Holyoke Center, Suite 727
1350 Massachusetts Avenue
Cambridge, MA 02138-5701

SIGNATURE: _____ DATE: _____

JOINT

PATENT APPLICATION

FOR

UNITED STATES LETTER PATENT

TO THE COMMISSIONER OF PATENTS AND TRADEMARKS:

BE IT KNOWN, that We, Kevin A. Jarrell, Matthew D. Shair have invented certain new useful improvements in COMBINATORIAL BIOLOGY of which the following is a specification:

Express Mail Label No: EK044517371US

Exchange 3076608 1

COMBINATORIAL BIOLOGY

The present application claims priority to United States Provisional Application USSN 60/114,909, filed January 5, 1999, the entire contents of which are incorporated herein by reference.

Background of the Invention

The ability of nature to produce small molecules having both structural complexity and biological potency has led to the development of countless number of therapeutic agents such as taxol, penicillin, and quinine, to name a few. Specific organisms found in nature have developed the ability to produce secondary metabolites in response to the challenges and needs that are encountered in their particular environment. In particular, through the random recombination and mutation of existing genetic material, the organism is able, in a combinatorial sense, to generate new biosynthetic enzymes that catalyze the assembly of new organic compounds (see, Verdine "The Combinatorial Chemistry of Nature" *Nature*, **1996**, *384*, 11).

Clearly, the ability to generate a diverse array of compounds reminiscent of these natural products would be desirable to increase the arsenal of available small molecules available for testing and use as therapeutic agents. Just as many scientists involved in the discovery and isolation of new natural products sample organisms from many different environments such as coral reefs, deep-sea hydrothermal vents, and tropical rainforests in their quest for structural diversity, both biologists and chemists have been searching for new ways to achieve this structural diversity by manipulating genetic material directly, or by generating novel synthetic pathways, respectively.

For example, certain researchers have attempted to alter the catalytic capability of synthetic enzymes that naturally produce interesting biologically active compounds by making specific changes to genes encoding particular enzymes (see, for example, Cortes et al., Science 268:1487, 1995; Kao et al., J. Am. Chem. Soc. 117:9105, 1995; Donadio et al., Science 252:675, 1991; WO 93/1363; U.S. Patent Number 5,824,513; WO 98/49315; U.S. Patent Number 5,652,116; U.S. Patent Number 5,824,774; WO 98/51695; U.S. Patent number 5,795,738; and

WO 98/01546). The hope is that, by modifying the gene encoding the synthetic enzyme, the researchers will be able to generate new enzymes with altered synthetic characteristics, which new enzymes will in turn generate new chemical compounds that are related to those produced by the naturally-occurring enzymes, and therefore are likely to have similar desirable biological activites.

Concurrently, several groups have also been interested in geneating combinatorial libraries of compounds synthesized using novel synthetic methods. Specifically, in many cases, researchers have developed "biased" libraries, in which all members share a particular characteristic, such as an ability to interact with a particular target ligand, or a characteristic structural feature designed to mimic a particular aspect of a class of natural compounds. For example, a number of libraries have been designed to mimic one or more features of natural peptides. Such "peptidomimetic" libraries include phthalimido libraries (WO 97/22594), thiophene libraries (WO 97/40034), benzodiazopene libraries (US 5, 288, 514), libraries formed by the sequential reaction of dienes (WO 96/03424), thiazolidinone libraries, libraries of metathiazanones and their derivatives (US 5, 549, 974), and azatide libraries (WO 97/35199) (for review of peptidomimetic technologies, see Gante, J., Angew. Chem. Int. Ed. Engl. 1994, 33, 1699-1720 and references cited therein). Each of these libraries has provided solid phase synthetic strategies for compounds possessing specific core functionalities, but none achieves the complexity of structure found in natural products, or in other lead compounds prepared through traditional chemical synthetic routes. Complex natural products commonly contain several different functionalities and often are rich in stereochemical complexity. Such diversity and complexity is difficult to achieve if the synthesis is restricted to compounds containing a specific core functionality.

Clearly, there remains a need to develop an efficient and powerful system for the generation of large numbers of compounds having unprecedented stereochemical, structural, topological and functional diversity. The present invention provides a unique method in which both the tools provided by nature and modern organic synthesis can be utilized.

Summary of the Invention

The present invention provides a method for merging combinatorial biosynthesis incorporated above with techniques of synthetic organic chemistry. In general, this method, combinatorial biology, involves 1) providing "starter units", wherein the starter units are capable of being accepted by the modular biosynthetic enzymatic machinery, and have incorporated therein a "functional handle" capable of reacting with specific functionality present on a solid support; 2) feeding these "starter units" into the modular biosynthetic enzymatic machinery, in vivo or in vitro, to obtain complex template molecules; and 3) further functionalizing the complex template molecules using synthetic organic chemistry to provide a collection of complex "unnatural" natural products having structural, topological, stereochemical and functional diversity.

In a preferred embodiment, the starter units are attached to solid support units prior to feeding the starter units into the modular biosynthetic enzymatic machinery, and thus the support bound starter units are fed into the biosynthetic enzymatic machinery to generate a collection of complex template structures. These template structures thus generated can then be further functionalized using synthetic organic chemistry, or can be further functionalized using any combination of synthetic organic chemistry and reintroduction into the biosythetic enzymatic machinery.

In another preferred embodiment, the starter units are fed to the modular biosynthetic enzymatic machinery prior to being attached to the solid support, and thus template structures are generated. In particularly preferred embodiments specific functionalities can also be incorporated into the template structures via the original starter unit capable of being recognized by an antibody and then purified. Alternatively or additionally, these templates can then be attached to solid support units and further functionalized using synthetic organic chemistry or any combination of synthetic organic chemistry and the biosynthetic enzymatic machinery.

Description of the Drawing

Figure 1 depicts the biosynthesis of several erethromycin derivatives using several different starter units.

Figure 2 depicts potential starter units with functional handles classified by the biosynthetic pathway.

Figure 3 depicts the functionalization of an iodoaryl compound to generate an arylacetylene via Sonogashiro/Castro-Stephens coupling reaction.

Figure 4 depicts certain preferred embodiments of the method of the present invention.

Figure 5 depicts specific complex template structures capable of further diversification using synthetic organic chemistry or biosynthetic pathways.

Detailed Description of the Invention

Recognizing the desirability of utilizing both the efficient and powerful methods of natural products biosynthesis and the the diverse repetoire of reactions available in synthetic organic chemistry, a method for merging combinatorial biosynthesis incorporated above with techniques of synthetic organic chemistry is provided. In general, this method, combinatorial biology, involves 1) providing "starter units", wherein the starter units are capable of being accepted by the modular biosynthetic enzymatic machinery, and have incorporated therein a "functional handle" capable of reacting with specific functionality present on a solid support; 2) feeding these "starter units" into the modular biosynthetic enzymatic machinery, in vivo or in vitro, to obtain complex template molecules; and 3) further functionalizing the complex template molecules using synthetic organic chemistry to provide a collection of complex "unnatural" natural products having structural, topological, stereochemical and functional diversity. As used herein, the term "starter unit" comprises any compound that can be incorporated into the biosynthetic pathway. For example, certain biosynthetic enzymes, such as polyketide synthases, utilize two different classes of "starter units", specifically "initiator" molecules and "extender" molecules, typically acetates or propionates. For the purposes of the present invention, either category, "initiator" or "extender" qualifies as a "starter" molecule.

As one of ordinary skill in the art will realize, because the starter units have incorporated therein a functional handle, they, or any of the products generated from these starter units, are capable of being attached to solid support units at any stage in the combinatorial biosynthetic pathway. In one preferred embodiment, the starter units are attached to solid support units prior

to feeding the starter units into the modular biosynthetic enzymatic machinery, and thus the support bound starter units are fed into the biosynthetic enzymatic machinery to generate a collection of complex template structures. These template structures thus generated can then be further functionalized using synthetic organic chemistry, or can be further functionalized using any combination of synthetic organic chemistry and reintroduction into the biosythetic enzymatic machinery. In another preferred embodiment, the starter units are fed to the modular biosynthetic enzymatic machinery prior to being attached to the solid support, and thus template structures are generated. In particularly preferred embodiments specific functionalities can also be incorporated into the template structures via the original starter unit capable of being recognized by an antibody and then purified. Alternatively or additionally, these templates can then be attached to solid support units and further functionalized using synthetic organic chemistry or any combination of synthetic organic chemistry and the biosynthetic enzymatic machinery.

Thus, the present invention represents a broadening of the concept of combinatorial biosynthesis to incorporate the advantages of organic synthetic techniques on the solid phase to generate increasingly complex "unnatural" natural products. Various characteristics of the starter units and the reactions utilized in preferred embodiments of the present invention are disucssed in more detail below; certain examples of the method of the present invention are also presented.

Biosynthetic Enzymatic Machinery

In principle, the inventive combinatorial biology methods may be applied to any biosynthetic pathway in which the synthetic enzymes will accept inventive starter molecules. In certain embodiments of the invention, the starter molecules are provided to living cells in which the synthetic enzymes are operating, and synthetic reactions in accordance with the present invention are performed in vivo. Alternatively, the biosynthetic pathway may be reproduced in vitro, and the starter molecules may be provided to the synthetic enzymes in that context. Preferred biosynthetic pathways to which the inventive technology may be applied include the animal fatty acid synthase pathway, the polyketide synthase pathway, the peptide synthetase

pathway, and the terpene (or isoprenoid) synthase pathway. In certain preferred embodiments of the invention, the naturally-occurring synthetic enzymes are employed and are simply provided with non-natural starter molecules as described herein.

Alternatively, as mentioned above, various researchers have made modifications to certain preferred biosynthetic enzymes that alter their catalytic properties (see, for example, Cortes et al., *Science* 268:1487, 1995; Kao et al., *J. Am. Chem. Soc.* 117:9105, 1995; Donadio et al., *Science* 252:675, 1991; WO 93/1363; U.S. Patent Number 5,824,513; WO 98/49315; U.S. Patent Number 5,652,116; U.S. Patent Number 5,824,774; WO 98/51695; U.S. Patent number 5,795,738; and WO 98/01546). Moreover, United States Patent Application Serial Number _____, entitled "Improved DNA Cloning", filed on even date herewith and incorporated herein by reference, describes a powerful system for the production of modified versions of biosynthetic enzymes, and in particular for the production of libraries of modified enzymes. Preferred embodiments of the present invention utilize such modified enzymes, and preferably libraries of modified enzymes, to catalyze synthetic reactions with inventive starter molecules.

To mention but one particularly preferred embodiment, the system described in the above-mentioned "Improved DNA Cloning" patent application can be utilized to generate a library of class I polyketide synthase enzymes in which a particular "AT" domain responsible for choosing an inventive altered starter molecule has been shuffled to a variety of different positions in the molecule. Exposure of such an enzyme library to the appropriate collection of natural and inventive starter and extender molecules will result in incorporation of inventive molecules at various locations in the polyketide-type compound being synthesized, so that a wide variety of different polyketide-type molecules can be produced by subsequent combinatorialization of the compounds generated by the synthetic enzyme.

Starter Units

As discussed above, the method of the present invention utilizes any combination of enzymatic machinery that can be employed for specific classes of biosynthetic reaction pathways. In determining the specific "starter units" that will be utilized in the synthesis,

consideration must also be made to the desired biosynthetic pathway to employ, and thus a desired family of compounds to be synthesized. For example, specific biosynthetic pathways such as polyketide synthases and peptide synthases, utilize particular starter units known to be accepted by the modular enzymes to produce a variety of natural products. In but one example, carboxylic acid building blocks having differing functional moieties such as methyl, ethyl, propyl groups are utilized by modular polyketide synthases to yield topologically, functionally and stereochemically diverse structures such as 6-deoxyerythronolide B and tylactone (see, Khosla, *Chem. Rev.* **1997**, *97*, 2577). Similarly, as shown in Figure 1, functionalized thioesters are utilized to generate a series of derivatives having structures similar to erythromycin (see, Khosla, *Chem. Rev.* **1997**, *97*, 2577). As one of ordinary skill in the art will realize, the modular enzymes that are produced by the shuffling may not result in the ability to predict the specific starter units that will be incorporated into the enzymatic machinery. Thus, the ability of a desired or random set of starter units to become attached to the solid support and subsequently diversified using combinatorial techniques (split-pool synthesis) becomes particularly important in identifiying *functional* starter units.

Clearly, as mentioned, each of the starter units provided must be capable of being accepted by the enzymes "machinery", and thus the functionality and structural topology must be compatible with the biosynthetic pathway. In but one example, as shown in Figure 2, it would be desirable for a terpene-based pathway to utilize functionalized derivatives of a farnasyl pyrophospate analog. Additionally, for an amino-acid based pathway (peptide synthases), it would be desirable to utilize functionalized derivatives of amino acids. Figure 2 additionally provides a collection of suitable "starter units" classified according to their utility in a specific biosynthetic pathway. Furthermore, these units can be derivatized and combinatorialized at a stage prior to feeding the units into the enzymatic machinery to produce structures having more diverse functionalities, or to produce structures having more complex topology. Thus, any structure may be utilized as a "starter unit", regardless of the complexity and functional diversity of the compound, provided that the starter unit is capable of being utilized by the enzymatic pathway.

In addition to the selection of suitable starter units that will ultimately be accepted by the enzymatic machinery, it is also necessary to ensure that each of the "starter units" that are fed into the modular biosynthetic enzymatic "machinery" preferably contains a suitable functional "handle" capable of attachment to functionality present on the solid phase. Figure 2 also depicts exemplary starter units having alkynes incorporated therein as functional handles. This can be effected by modifying the starter units that will be utilized in a particular biosynthetic pathway to incorporate the specific desired functionality for coupling to the solid support. As one of ordinary skill in the art will realize, attachment to the solid phase can occur prior to feeding the starter units into the enzymatic machinery, or after feeding the starter units into the enzymatic machinery to generate complex template structures to be attached to the solid support. Because attachment can occur prior to or after exposure to the enzymatic machinery, it is particularly preferred that the functional handle is chemically robust and thus is capable of withstanding any of the reaction conditions encountered in the enzymatic machinery. For this reason, particularly robust functionalities such as alkynes, olefins and iodoalkenes, are particularly preferred, although the method of the present invention is not limited to these functionalities.

In an exemplary embodiment, the starter units utilized in the present invention have alkyne functionalities incorporated therein as handles for attachment to the solid support. To illustrate the incorporation of "handles" into the starter units, Figure 3 depicts the use of the Sonagashiro/Castro-Stephens reaction, in which unactivated aryl halides react with copper acetylides to give good yields of arylacetylenes, for the conversion of an aryl iodide to an aryl acetylene in the iodo substituted aryl thioester derivative.

An additional important matter to take into consideration when utilizing a starter unit having "handles", is the number of "handles" that will ultimately be incorporated into the complex template as a result of the biosynthesis. For example, if a simpler (less structurally complex) starter unit is utilized (for example one that is used in successive condensation reactions, as exemplified in the polyketide synthases) a larger number of handles will be incorporated. Thus, if one were to incorporate a handle into a starter unit that is utilized in successive "rounds" by the enzymatic machinery, a template structure would result having many handles incorporated therein. Thus, to overcome this problem, it is preferable to 1) utilize starter

units that are only incorporated once, as depicted in Figures 2 and 3 by the use of the aryl functionality, and as may be dictated by the selectivity of specific reactions in the enzymatic pathway; 2) to incorporate the handle into a limited number of starter units (for example, one or two units) involved in the synthesis, or 3) to utilize more complex starter units, as depicted in Figure 2, for the polyketide pathway, and thus the enzymatic machinery does not need to utilize as many of the starter units and therefore fewer handles will be incorporated. As one of ordinary skill in the art will realize, when selecting specific starter it is necessary to take into consideration the specific enzymatic biosynthetic pathway to be utilized (and the specificity of incorporation of certain starters) and the number of "handles" desired in the resulting template structure. Once the "handles" have been incorporated into the "starter units", the starter units can either be fed directly into the enzymatic machinery or coupled directly to the solid support.

Whether the starter unit or the template structure is utilized for coupling to the solid support, (which, as one of ordinary skill in the art will realize, also has an alkyne, or other desirable functionality, bound thereto via a cleavable bond utilizng standard synthetic organic techniques) the two alkynes incorporated can then be coupled, in a preferred embodiment, using a Glaser Coupling reaction. Specifially, in this reaction, the addition of copper (II) acetate to the reaction medium effects coupling of the two components to yield a solid support having bound thereto a starter unit, or a template structure, via a diyne functionality. In other exemplary embodiments, suitable functional handles which can also be easily incorporated into starter units and solid support units using starndard techniques of synthetic organic chemistry, include olefins and iodoalkenes. Thus, coupling of the components can be effected using olefin metathesis and Stille Coupling reactions, respectively. One of ordinary skill in the art will realize that although the abovementioned functionalities are particularly preferred, other functionalities that will not interfere with, or be altered by, the chemistry being employed by the enzymatic modular machinery may also be utilized.

Combinatorial Bioorganic Synthesis

Once the selection of desired starting materials is achieved and functional handles are incorporated therein for attachment to the solid phase and optionally for purification purposes,

several variations of the present invention may be employed to obtain "unnatural" natural products. Figure 4 refers to preferred embodiments of the present invention and a discussion of each of these embodiments with reference to Figure 4 is presented below.

As shown in Figure 4, genetic manipulation of the modular enzymatic machinery, yields a mutated biosynthetic pathway to be utilized for the generation of novel "unnatural" natural products. Once one or more desired mutated biosynthetic pathways are developed, either randomly or with a particular target in mind, as discussed above, the tools of synthetic organic chemistry can be employed prior to, concurrently with, or after exposure to the enzymatic machinery to further diversify and increase the universe of "unnatural" natural products that can be generated. Specifically, referring to Figure 4, selected starter units can be fed into this mutated biosynthetic pathway in two different ways according to different embodiments of the present invention.

In one particularly preferred embodiment, as shown by pathway A, the selected starter units having the functional handles incorporated therein can be fed directly into the mutated biosynthetic enzymatic machinery. Thus, after exposure to the enzymatic machinery, a unique template compound can be obtained and at this stage can be purified, preferably using an antibody recognition element, or can also be attached to a solid support unit. Subsequently, after attachment to a solid support unit, the template could be reintroduced into the enzymatic machinery (same or different) to thus obtain a modified template. Alternatively, and/or additionally, the template could be utilized in split-pool organic synthesis to generate combinatorial libraries of complex "unnatrual" natural products.

In another particularly preferred embodiment, as shown in pathway B, the selected starter units having the functional handles incorporated therein can be attached to the solid support prior to being exposed to the biosynthetic enzymatic machinery. Alternatively, or additionally, a collection of starter units could be generated using combinatorial synthetic organic chemistry. After one or more starter units are fed into the enzymatic machinery, the starter units are transformed into a support bound template structure. As Figure 4 depicts, the template structure resulting from an initial exposure to a single biosynthetic pathway could be reexposed to the same biosynthetic or to another biosynthetic pathway to further modify the structure.

Alternatively, or additionally, the solid support template structure could be diversified using split and pool techniques and synthetic organic chemistry. As one of ordinary skill in the art will realize, any combination of the enzymatic machinery and split-pool synthetic organic techniques can be employed to increase the structural, functional, topological and stereochemical diversity to generate a collection of unique "unnatural" natural products.

Referring to the embodiments discussed above, either split-pool or parallel synthesis methods can be employed at each stage of the inventive method to provide a desired collection of compounds. For example, if the starter unit is not initially attached to the solid support a parallel synthesis technique is preferably utilized so that the product resulting from the enzymatic machinery can be identifed using spatial encoding methods. Additionally, subsequent identification of compounds using standard methods such as nuclear magnetic resonance spectroscopy or mass spectrometry can also be employed to identify specific compounds. Depending on the size of the library of compounds desired for the synthesis, the idenification of individial compounds may be prohibitively time consuming and therefore a spatial encoding method may be more particularly preferred.

As discussed above, in one embodiment of the invention, the template structures are generated in solution, rather than on a solid support unit. In a preferred embodiment, for the generation of a collection of compounds in solution, a parallel synthesis technique is utilized, in which all of the products are assembled separately in their own reaction vessels. In a particularly preferred parallel synthesis procedure, a microtitre plate containing n rows and m columns of tiny wells which are capable of holding a few milliliters of the solvent in which the reaction will occur, is utilized. One of ordinary skill in the art will realize that this particular procedure is most useful when smaller libraries are desired, and the specific wells can provide a ready means to identify the library members in a particular well.

In another more particularly preferred embodiment of the present invention, a solid phase synthesis technique is utilized for the biosynthesis of the templates and the diversified structures, or alternatively for the synthesis of the diversified structures produced from template structures generated in solution as described above. As discussed in detail, the starter units or the template structures are attached to the solid phase directly or though a linking unit, depending on the stage

of the procedure a solid phase synthesis is desired. Advantages of solid phase techniques, most particularly at the stage when the template structures are functionalized using synthetic organic techniques, include the ability to more easily conduct multi-step reactions and the ability to drive reactions to completion because excess reagents can be utilized and the unreacted reagent washed away. Perhaps one of the most significant advantages of solid phase synthesis is the ability to use a technique called "split and pool", in addition to the parallel synthesis technique, develped by Furka. (Furka et al., *Abstr. 14th Int. Congr. Biochem.*, Prague, Czechoslovakia, **1988**, *5*, 47; Furka et al., *Int. J. Pept. Protein Res.* **1991**, *37*, 487; Sebestyen et al., *Bioorg. Med. Chem. Lett.*, **1993**, *3*, 413.) In this technique, a mixture of related compounds can be made in the same reaction vessel, thus substantially reducing the number of containers required for the synthesis of very large libraries, such as those containing as many as or more than one million library members. As an example, the solid support templates or starter units can be divided into n vessels, where n represents the number species of reagent A, or the number of different biosynthetic pathways created by the shuffling procedure, to be reacted with the template structures or starter units. After reaction, the contents from n vessels are combined and then split into m vessels, where m represents the number of species of reagent B, or the number of different biosynthetic pathways created by the shuffling procedure, to be reacted with the scaffold structures. This procedure is repeated until a desired collection of structures is obtained to yield a library of "unnatural" natural products.

The use of solid phase techniques in the present invention may also include the use of a specific encoding technique. Specific encoding techniques have been reviewed by Czarnik. (Czarnik, A.W., *Current Opinion in Chemical Biology*, **1997**, *1*, 60.) As used in the present invention, an encoding technique involves the use of a particular "identifiying agent" attached to the solid support, which enables the determination of the structure of a specific library member without reference to its spatial coordinates. One of ordinary skill in the art will also realize that if smaller solid phase libraries are generated in specific reaction wells, such as 96 well plates, or on plastic pins, the reaction history of these library members may also be identified by their spatial coordinates in the particular plate, and thus are spatially encoded. It is most preferred,

however for large combinatorial libraries, to use an alternative encoding technique to record the specific reaction history.

Examples of particulary preferred alternative encoding techniques that can be utilized in the present invention include, but are not limited to, spatial encoding techniques, graphical encoding techniques, including the "tea bag" method, chemical encoding methods, and spectrophotometric encoding methods. Spatial encoding refers to recording a reaction's history based on its location. Graphical encoding techniques involve the coding of each synthesis platform to permit the generation of a relational database. Examples of preferred spectrophotometic encoding methods include the use of mass spectroscopy, fluorescence emission, and nuclear magnetic resonance spectroscopy. In a most preferred embodiment, chemical encoding methods are utilized, which uses the structure of the reaction product to code for its identity. Decoding using this method can be performed on the solid phase or off of the solid phase. One of ordinary skill in the art will realize that the particular encoding method to be used in the present invention must be selected based upon the number of library members desired, and the reaction chemistry employed.

Subsequent characterization of the library members, which can include either the scaffolds obtained after the biosynthetic pathway or the complex molecules obtained after diversification using synthetic organic chemistry, can be performed using standard analytical techniques, such as mass spectrometry, Nuclear Magnetic Resonance Spectroscopy, and gas chromatrograpy. One of ordinary skill in the art will realize that the selection of a particular analytical technique will depend upon whether the inventive library members are in the solution phase or on the solid phase.

Reactions at latent functionality in the inventive templates to generate "nonnatural" compounds

Once the inventive templates have been synthesized as discussed above, diversification reactions may be employed at each of the different latent functionality sites present in the template structures. As mentioned previously, these functionalized templates may then also be reintroduced into the enzymatic machinery for further functionalization and structural changes, as well as be further functionalized using synthetic organic chemistry, until a desired collection

of compounds is obtained. One of ordinary skill in the art will appreciate that at any stage the reactivity of a particular functionality present in the template structure must be considered when selecting particular reagents for diversification.

Figure 5 depicts specific natural products which are capable of being generated using the method of the present invention, or alternatively derivatives of which are also capable of being generated using the method of the present invention. Specifically, these structures exemplify the variety of sites of latent functionality at which diversification can be achieved using synthetic combinatorial techniques. As discussed previously, this can be achieved using solution or solid phase methods, using parallel or split-pool techniques. It is particularly preferred, however, to utilized solid phase split-pool techniques. Examples of specific reactions to which some or all of the systems depicted in Figure 5 can be subjected to in solution or on the solid support include, but are not limited to, i) addition of nucleophiles (such as primary and secondary amines), ii) functionalization of free hydroxyls with electrophiles (for examples isocyanates, anhydrides, or acid chlorides, iii) opening of epoxides with nucleophiles, such as amines, under ytterbium catalysis, iv) functionalization of aromatic rings, specifically functionalization at an aryl iodide by conversion to such structures as amines, amides, aromatic rings, alkenes, alkynes, and heterocycles using palladium catalyzed chemistry such as Buchwald-Hartwig aminations, Heck and Stille couplings, Sonogashira/Castro-Stephens couplings, Suzuki and Stille couplings, and carbonylations. Furthermore, resulting aryl alkynes can undergo rhodium-catalyzed hydroacylation and azide cycloaddition and nitrone and nitrile oxide cycloadditions. Other examples of diversification reactions at potential sites include reactions at amines and amides. For example, amides may be functionalized using a Mitsunobu reaction to generate alcohols such as straight chain, branched, and cyclic alcohols.

Additionally, for each of the compounds produced by the inventive method, further reactions may be employed to attach biomolecules (such as polysaccharides, nucleic acids, or peptides) or polymers to appropriate functionalities.

One of ordinary skill in the art will realize that the above examples are representative of the reactions that can be used to diversify not only the templates, but also the starter units, of the present invention and are not intended to be exclusive. Rather, all equivalents thereof are

intended to be within the scope of the presently claimed invention. A skilled artisan in the field of synthetic organic chemistry will be able to readily identify those reagents capable of reacting to create further diversity at selected sites in the inventive template structures and starter units to generate compounds and libraries of compounds reminiscent of natural products. The inventive method is particularly useful for the generation of such compounds because it incorporates the efficiency and creativity of a method for manipulating the enzymatic reactions that produce such complex structures in nature with the arsenal of reactions available using synthetic organic techniques. For example, the generation of complex templates can be difficult utilizing only synthetic organic techniques, however, nature effeciently and elegantly provides these templates. Combining this with synthetic organic chemistry enables the use of reactions, such as palladium catalyzed reactions that are not available in nature and thus the best of both systems can be utilized to generate compounds having unprecedented structural, topological, stereochemical and functional diversity.

Uses

The methods, compounds and libraries generated by the method of the present invention can be utilized in various disciplines. For example, the complex molecules generated in the method of the present invention may modulate the biological activity of a biological target, such as a protein, nucleic acid, lipid or combination thereof. In one preferred embodiment, the compounds generated by the method of the invention are utilized in chemical genetics assays to alter, i.e. inhibit or initiate, the action of such biological molecules. Alternatively or additionally, the compounds may be used in in vitro assays, or any other system that allows detection of a chemical or biological function.

In a particularly preferred embodiment of the invention, one or more inventive compounds is contacted with a biological target having a detectable biochemical activity. Such biological targets include, for example, enzymes, receptors, subunits involved in the formation of multimeric complexes. Such multimeric complex subunits may be characterized by catalytic capabilities (such as, for example, an ability to catalyze substrate conversion), or may alternatively be primarily active in binding to one or more other molecule. The biological target

can be provided in the form of a purified or semi-purified composition, a cell lysate, a whole cell or tissue, or even a whole organism. The level of biochemical activity is detected in the presence of the compound, and a statistically significant change in the biochemical activity, relative to the level of biochemical activity in the absence of the compound, identifies the compound as a modulator, e.g. inhibitor or potentiator of the biological activity of the target protein. In some cases, particularly where assays are done on whole cells or organisms, the effect of the chemical compound may be to alter the amount, in addition to or instead of the activity, of the particular biological target. "Modulators", therefore, are chemical compounds that alter the level or activity of a particular target molecule.

In one particularly preferred embodiment of the present invention, multiple compounds are assayed simultaneously in a high-throughput format, preferably allowing simultaneous analysis of at least 500,000 compounds, preferably at least 1,000,000 compounds, and most preferably at least or more than 2,000,000 compounds.

As discussed above, once a specific desired effect on a biological target has been associated with a particular compound of the inventive library, the compounds of the present invention may be utilized as a therapeutic agent for a particular medical condition. A therapeutic agent for use in the present invention may include any pharmacologically active substances that produce a local or systemic effect in animals, preferably mammals, or humans. The term thus means any substance intended for use in the diagnosis, cure, mitigation, treatment or prevention of disease or in the enhancement of desirable physical or mental development and conditions in an animal or human. The therapeutic agent may be administered orally, topically or via injection by itself, or additionally may be provided as a pharmaceutical composition comprising the therapeutic agent and a biologically acceptable carrier. The inventive compositions can be, but are not limited to an aqueous solutions, emulsions, creams, ointments, suspensions, gels, and liposomal suspensions. Particularly preferred biologically acceptable carriers include but are not limited to water, saline, Ringer's solution, dextrose solution and solutions of ethanol, glucose, sucrose, dextran, mannose, mannitol, sorbitol, polyethylene glycol (PEG), phosphate, acetate, gelatin, collagen, Carbopol, and vegetable oils. It is also possible to include suitable preservatives, stabilizers, antioxidants, antimicrobials, and

buffering agents, for example including but not limited to BHA, BHT, citric acid, ascorbic acid, and tetracycline. The therapeutic agents of the presently claimed invention may also be incorporated or encapsulated in a suitable polymer matrix or membrane, thus providing a sustained-release delivery device suitable for implantation near the site to be treated locally.

As one of ordinary skill in the art will realize, the amount of the therapeutic agent required to treat any particular disorder will of course vary depending upon the nature and severity of the disorder, the age and condition of the subject, and other factors readily determined by one or ordinary skill in the art.

In alternative embodiments, the compounds and libraries of the present invention may also be used for the development of cosmetics, food additives, pesticides, and lubricants to name a few. Furthermore, the compounds and libraries of the present invention may also be used for the development of novel catalysts and materials. For example, the inventive compounds may be useful as ligands for transition metal catalysts and the inventive libraries may be useful for the rapid identification of novel ligands. These compounds and libraries of compounds may also function by acting in concert with a particular transition metal catalyst to effect a particular desired chemical reaction. Additionally, the inventive compounds and libraries of compounds are also useful in the area of materials science. Because of the reactive moieties present in these compounds, molecules such as lipids and other polymeric materials may be attached and thus generate potentially important biomaterials.

One of ordinary skill in the art will realize that the present invention is not intended to be limited to the abovementioned uses, but rather may be employed in many contexts and disciplines.

Claims

What we claim is:

1. A method for the combinatorial biosynthesis of one or more compounds comprising:

a) providing one or more starter units, wherein said one or more starter units have incorporated therein a functional handle capable of reacting with a functionality present on a solid support unit;

b) attaching said one or more starter units to a solid support unit;

c) feeding said one or more support bound starter units into one or more biosynthetic enzymatic machinery systems to generate a collection of template structures;

d) functionalizing said template structures using synthetic organic chemistry; and

e) repeating steps c) and/or d) until a desired collection of structures is generated.

2. The method of claim 1, further comprising functionalizing said collection of structures generated in step e) to provide a collection of unnatural natural products.

3. A method for the combinatorial biosynthesis of one or more compounds comprising:

a) providing one or more starter units, wherein said one or more starter units have incorporated therein a functional handle capable of reacting with a functionality present on a solid support unit;

b) attaching said one or more starter units to one or more solid support units;

c) feeding said one or more starter units into one or more biosynthetic enzymatic machinery systems to generate a collection of template structures;

d) functionalizing said collection of template structures to generate a collection of unnatural natural products.

4. A method for the combinatorial biosynthesis of one or more compounds comprising:

a) providing one or more starter units, wherein said one or more starter units have incorporated therein a functional handle capable of reacting with a functionality present on a solid support unit;

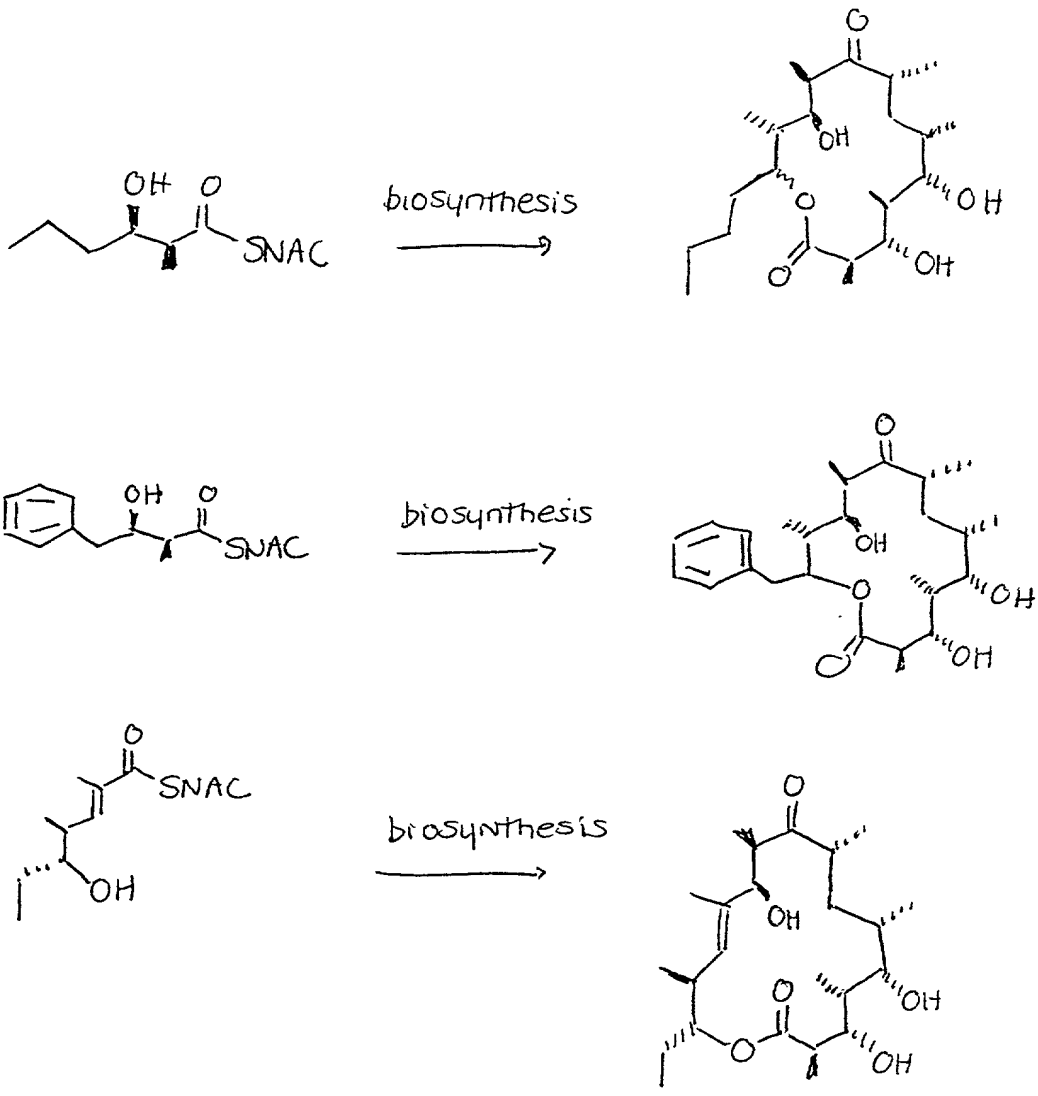
b) feeding said one or more starter units into one or more enzymatic machinery sytstems to generate a collection of template structures;

c) attaching said collection of template structures to solid support units; and

d) functionalizing said collection of support bound templates to generate a collection of unnatural natural products.

Abstract

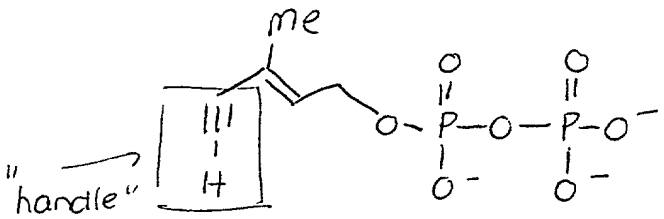
The present invention provides a method for merging combinatorial biosynthesis with techniques of synthetic organic chemistry. In general, this method, combinatorial biology, involves 1) providing "starter units", wherein the starter units are capable of being accepted by the modular biosynthetic enzymatic machinery, and have incorporated therein a "functional handle" capable of reacting with specific functionality present on a solid support; 2) feeding these "starter units" into the modular biosynthetic enzymatic machinery, in vivo or in vitro, to obtain complex template molecules; and 3) further functionalizing the complex template molecules using synthetic organic chemistry to provide a collection of complex "unnatural" natural products having structural, topological, stereochemical and functional diversity. In one preferred embodiment, the starter units are attached to solid support units prior to being exposed to the modular biosynthetic enzymatic machinery. In another preferred embodiment, the starter units are exposed to the modular biosynthetic enzymatic machinery directly and thus the templates produced by the machinery are thus attached to the solid support. The starter units and the template structures used in the method of the present invention can both be diversified using synthetic organic chemistry.



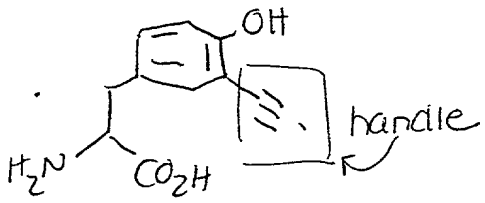
(Khosla, Chem. Rev., 1997, 97, 2587.) Figure 1

Potential starter units (w/ functional handles)

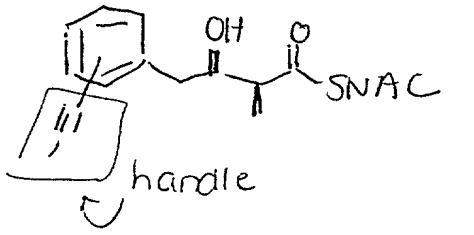
① Terpene-based pathway



② Amino-acid based pathway



③ Polyketide pathway



OR

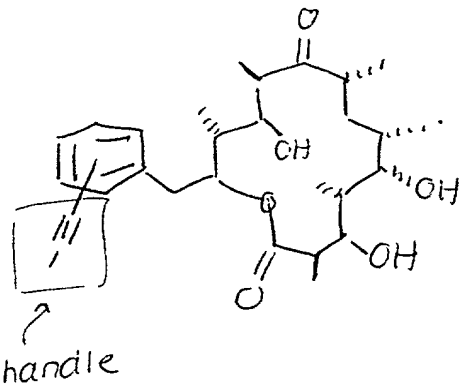
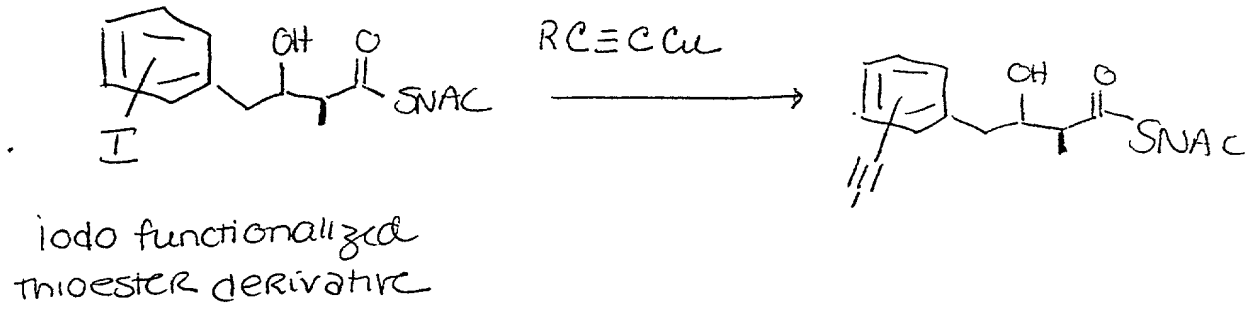


FIGURE 2



Iodo functionalized thioester derivative

Figure 3

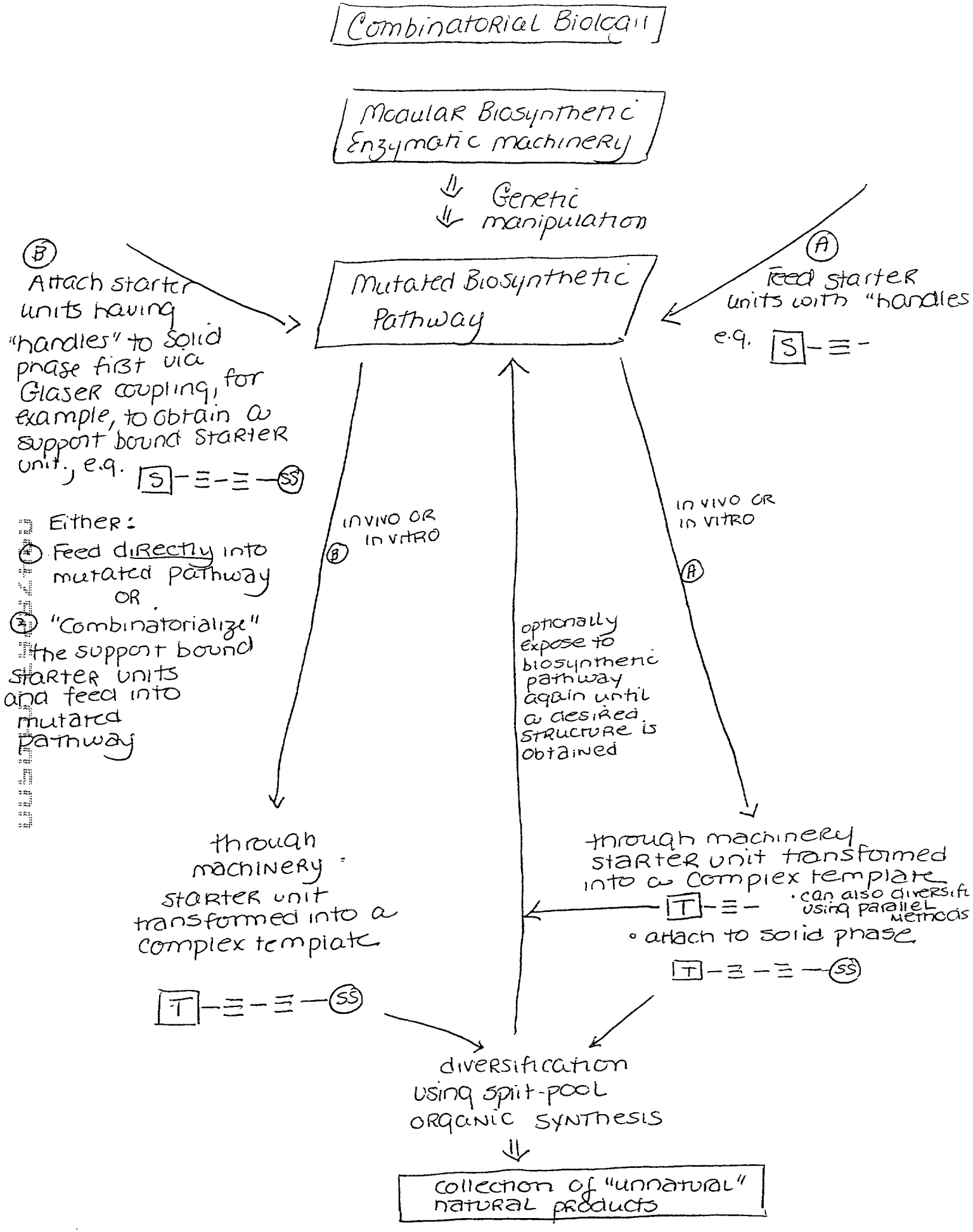
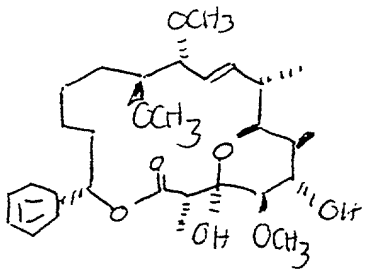
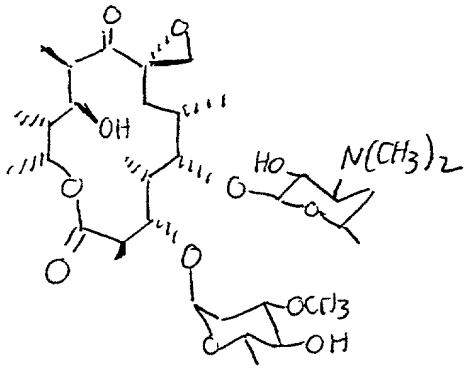

Combinatorial Biolcgy
Modular Biosynthetic Enzymatic machinery
Genetic manipulation
Mutated Biosynthetic Pathway
(B) Attach starter units having "handles" to solid phase first via Glaser coupling, for example, to obtain a support bound starter unit., e.g. S-≡-≡-SS
(A) Feed starter units with "handles" e.g. S-≡-
Either:
(A) Feed directly into mutated pathway
OR
(B) "Combinatorialize" the support bound starter units and feed into mutated pathway
IN VIVO OR IN VITRO
(B)
IN VIVO OR IN VITRO
(A)
optionally expose to biosynthetic pathway again until a desired structure is obtained
through machinery starter unit transformed into a complex template
T-≡-≡-SS
through machinery starter unit transformed into a complex template
T-≡-
• can also diversify using parallel methods
• attach to solid phase
T-≡-≡-SS
diversification using split-pool organic synthesis
collection of "unnatural" natural products

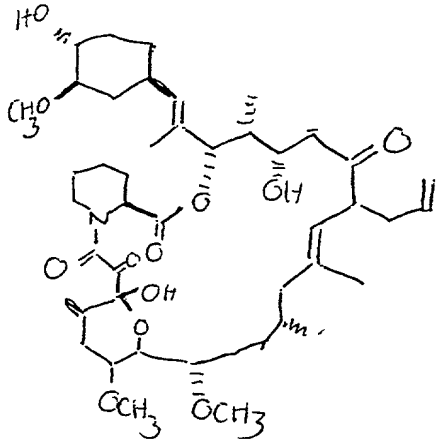
FIGURE 4



Soraphen(A)



Oleandomycin



FK 506

Figure 5

DECLARATION

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name,

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled:

COMBINATORIAL BIOLOGY

the specification of which (I authorize Choate, Hall & Stewart to check one of the following, three choices, and fill in the blanks, if applicable):

__X__ is attached hereto

____ was filed on _____ as Application
Serial No._____ and amended on _____ (if applicable).

_____ was filed as PCT international application No. _____,
on _____ and was amended under PCT Article 19
on_____ (if applicable).

I hereby state that I have reviewed and understood the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

I acknowledged the duty to disclose information which is material to the examination of this application in accordance with Title 37, Code of Federal Regulations, §1.56.

I hereby claim foreign priority benefits under Title 35, United States Code, §119 of any foreign application(s) for patent or inventor's certificate listed below and have also identified below any foreign application for patent or inventor's certificate having a filing date before that of the application on which priority is claimed:

Prior Foreign Application(s):			Priority Claimed	
(Number)	(Country)	(Day/Month/Year/Filed)	Yes	No
(Number)	(Country)	(Day/Month/Year/Filed)	Yes	No

I hereby claim the benefit under Title 35, United States Code, §120 of any United States application(s) or PCT international application(s) designating the United States of America

listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States application in the manner provided by the first paragraph of Title 35, United States Code, §112, I acknowledge the duty to disclose material information as defined in Title 37, Code of Federal Regulations, §1.56 which became available between the filing date of the prior application and the national or PCT international filing date of this application:

(Application Serial No.)	(filing date)	(status-patented, pending, abandoned)
(Application Serial No.)	(filing date)	(status-patented, pending, abandoned)

PCT Applications designating the United States:

(PCT Appl. No.)	(U.S.S.N.)	(status-patented, pending, abandoned)

I hereby claim the benefit under Title 35, United States Code, §119(e) of any United States provisional application(s) listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States application in the manner provided by the first paragraph of Title 35, United States Code, §112, I acknowledge the duty to disclose information which is material to patentability as defined in Title 37, Code of Federal Regulations, §1.56 which became available between the filing date of the prior application and the national filing date of this application.

Provisional Application(s):

Application Number	Filing Date	Status
60/114,909	**January 5, 1999**	**pending**

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United State Code and that such willful false statements may jeopardize the validity of the application or any patents issued thereon.

Full name of first inventor Kevin A. Jarrell

Inventor's signature _____ Date: _____

Residence 3 Acorn Lane, Lincoln, MA 01773

Citizenship _____

Post Office Address Same

Full name of second inventor Matthew D. Shair

Inventor's signature _____ Date: _____

Residence 16 Foskett Street, Apt. 2, Somerville, MA 02144

Citizenship _____

Post Office Address Same

ATTORNEY'S DOCKET NO.: 0342941-0037

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant: Jarrell, et al.
Serial Number:
Filed: January 5, 2000
For: COMBINATORIAL BIOLOGY

Assistant Commissioner for Patents
Washington, D.C. 20231

Sir:

APPOINTMENT OF ATTORNEY

The undersigned hereby appoints Brenda Herschbach Jarrell, Registration No. 39,223; Sam Pasternack, Registration No. 29,576; David J. Powsner, Registration No. 31,868; Kevin M. Tormey, Registration No. 41,351; Elizabeth Nugent, Registration No. 43,839, Valarie B. Rosen, Registration No. P-45,698 and Stanley C. Mah, Registration No. P-46,189 as its attorneys and agents for prosecution of matters relating to the above-identified patent application and to conduct all business in the United States Patent and Trademark Office.

All correspondence should be sent to Brenda Herschbach Jarrell, Choate, Hall & Stewart, Exchange Place, 53 State Street, Boston, Massachusetts 02109.

Respectfully submitted,

_____
Name:_____
Title:_____
On behalf of _____

Dated: _____, 2000